Article:

Flack, Tessa R, Andrews, Timothy J orcid.org/0000-0001-8255-9120, Hymers, Mark et al. (7 more authors) (2015) Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. Cortex; a journal devoted to the study of the nervous system and behavior. pp. 14-23. ISSN 1973-8102

https://doi.org/10.1016/j.cortex.2015.03.002

RESPONSES IN THE RIGHT POSTERIOR SUPERIOR TEMPORAL SULCUS SHOW A FEATURE-BASED RESPONSE TO FACIAL EXPRESSION

Tessa R. Flack, Timothy J. Andrews, Mark Hymers, Mohammed Al-Mosaiwi, Samuel P. Marsden, James W.A. Strachan, Chayanit Trakulpipat, Liang Wang, Tian Wu and Andrew W. Young.

Department of Psychology and York Neuroimaging Centre, University of York, Heslington, York, YO10 5DD, UK

*Corresponding author: tessa.flack@york.ac.uk, +44 1904 323 165

Figures: 4

Tables: 2

Abstract: 291

Introduction: 876

Discussion: 1138

Text: 7359

Key words: emotion, facial expression, holistic perception, posterior STS

ABSTRACT

The face-selective region of the right posterior superior temporal sulcus (pSTS) plays an important role in analysing facial expressions. However, it is less clear how facial expressions are represented in this region. In this study, we used the face composite effect to explore whether the pSTS contains a holistic or feature-based representation of facial expression. Aligned and misaligned composite images were created from the top and bottom halves of faces posing different expressions. In Experiment 1, participants performed a behavioural matching task in which they judged whether the top half of two images was the same or different. The ability to discriminate the top half of the face was affected by changes in the bottom half of the face when the images were aligned, but not when they were misaligned. This shows a holistic behavioural response to expression. In Experiment 2, we used fMR-adaptation to ask whether the pSTS has a corresponding holistic neural representation of expression. Aligned or misaligned images were presented in blocks that involved repeating the same image or in which the top or bottom half of the images changed. Increased neural responses were found in the right pSTS regardless of whether the change occurred in the top or bottom of the image, showing that changes in expression were detected across all parts of the face. However, in contrast to the behavioural data, the pattern did not differ between aligned and misaligned stimuli. This suggests that the pSTS does not encode facial expressions holistically. In contrast to the pSTS, a holistic pattern of response to facial expression was found in the right inferior frontal gyrus. Together, these results suggest that pSTS reflects an early stage in the processing of facial expression in which facial features are represented independently.

1. INTRODUCTION

Interpreting the facial expressions of others is important to effective social interaction (Bruce & Young, 2012). Facial expressions result from characteristic patterns of movement of the facial muscles that can easily be seen in static photographs (usually showing the apex of the movement itself) or in videos (Johnston et al., 2013). However, little is known about how expressions are encoded at the neural level. The most widely-used neural model of face perception (Haxby et al., 2000) proposes that the superior temporal sulcus (STS) is a key neural structure in the perceptual analysis of facial expressions, and this is borne out by a number of studies that have implicated STS in neural responses to expression (Calder & Young, 2005; Psalta, Young, Thompson & Andrews, 2014) and social perception from visual cues (Allison, Puce & McCarthy, 2000).

Relatively few studies address the question of *how* STS encodes expression. Said et al. (2010) were able to demonstrate that patterns of activation to different facial expressions across voxels in posterior STS (pSTS) were correlated with the rated perceptual similarities of the expressions themselves, suggesting that the functional organisation of pSTS reflects this underlying perceptual structure. Similarly, Harris et al. (2012) found that right pSTS responded to changes in facial expression regardless of whether or not these changes crossed or remained within emotional category boundaries, which again suggests a form of encoding that is largely driven by the perceptual input. Importantly, Harris et al. (2014) showed that right pSTS is relatively insensitive to contrast reversal, which implies that the critical perceptual input for pSTS involves feature shapes. Contrast reversal is known to have a dramatic effect on face identity recognition, but it has relatively little effect on the recognition of expression because information about feature shapes that is critical to interpreting facial expressions is conveyed through the position of edges that remain largely invariant to contrast reversal (Bruce & Young, 1998).

Here, we take the study of the perceptual representation used by pSTS a step further by asking whether it represents features such as the eyes and mouth independently from each other, or as part of a perceptual whole (the face). The critical test of holistic processing that we use for this purpose is the expression composite effect. Composite effects have been demonstrated in many studies of facial identity perception (Young, Hellawell & Hay, 1987; Rossion, 2013), but their extension to understanding facial expression perception is less well-known. The paradigm involves combining the top half of one facial expression with the bottom half of another expression and determining whether this combination of different parts results in the perception of a new whole expression (Prazak & Burgund, 2014, Calder & Jansen, 2005, Calder et al., 2000, Palermo et al., 2011). The critical test of holistic perception involves contrasting performance between images in which the top and bottom halves are aligned into a highly face-like overall configuration, or misaligned so that they are less face-like. Contrasting aligned and misaligned versions of composite images created from the top and bottom parts of different facial expressions makes it possible to differentiate responses based on face features, which will be equivalent across aligned and misaligned image variants, from holistic responses that will only be evident for aligned and not for misaligned images.

In this study, we used the facial expression composite effect to investigate whether neural responses to facial expression in right pSTS reflect feature changes or are dependent on the face as a perceptual whole. To do this, we first established in a behavioural study that the stimuli and presentation parameters we intended to use in fMRI elicited a robust expression composite effect. We then compared neural responses in right pSTS to composite expressions in which the top (eye region) and bottom (mouth region) parts were aligned into an overall face-like configuration with neural responses to misaligned stimuli created by shifting one part horizontally with respect to the other (see Figure 1). Misalignment still allows the separated parts of the face to be encoded as features, but it interferes with the integration of expressive information from the eye and mouth region into a perceptual whole (Calder et al., 2000).

Our fMRI experiment used a block design adaptation paradigm in which participants viewed blocks comprising a series of facial expressions that were all the same (no change condition) or that varied across the top half of each image (top change condition) or across the bottom half of each image (bottom change condition). During these blocks, participants were asked to fixate between the eyes (i.e. in the top half of each face) and further to encourage fixation they had to detect the presentation of an occasional small red spot at the fixation point. The no change condition, with identical stimuli throughout the block, served as a baseline that will lead to maximal adaptation of neural responses, and the top change or bottom change conditions measured any release from adaptation in neural regions that can encode these changes. The stimuli were aligned into overall face-like composites, or horizontally misaligned so that they were not face-like (see Figure 1), allowing us to establish whether the pattern of neural responses across conditions involving no change, top change, or bottom change was dependent on the presence of a face-like (aligned) configuration.

2. MATERIAL AND METHODS

2.1. *Participants*

Sixteen participants took part in experiment 1 (8 male, 8 female, mean age 27.6 ± 4.4). Twenty-seven participants took part in experiment 2 (17 male, 10 female, mean age 24.7 ± 5.0). All participants had normal or corrected-to-normal vision, with no known history of neurological disorder and no abnormalities that were immediately evident from structural MRI in experiment 2. Written consent was obtained from all participants and the studies were approved by the York Neuroimaging Centre Research Ethics Committee and the Department of Psychology Ethics Committee at the University of York. One participant was removed from the fMRI analysis due to excessive head movement.

2.2. *Experiment 1*

2.2.1. *Stimuli and Design*

The initial behavioural study used to validate key procedural parameters, experiment 1, involved six conditions. Stimuli consisted of aligned composite and misaligned non-composite images of greyscale faces which either had the same top and bottom half (no change), the same bottom half with the top half varying in expression (top change), or the same top half with the bottom half varying in expression (bottom change). The top and bottom half images were separated by a gap of 5 pixels, in line with the procedural strictures of Rossion (2013). Examples can be seen in Figure 1. Top and bottom half face images were derived from Ekman faces taken from the FEEST set (Young et al. 2002). Two individuals posing four facial expressions (fear, anger, happiness and disgust) were used to create the stimuli. These individuals were selected on the basis of a high recognition rate for all expressions and consistency of the action units used to pose each expression (Young et al. 2002).

Aligned or misaligned images were presented in sequential pairs in which both members of the pair had aligned constituent parts or both had misaligned parts. In misaligned pairs the offset was to the left in half the trials, or to the right in the other half. Images were presented using an LCD monitor, approximately 57 cm from the participant. The images were presented for 750 ms each, with a 750 ms inter-stimulus interval. Participants were instructed to only look at the top half of the face. There was a fixation cross located between the eyes on each ISI and a chin rest was used to help participants maintain fixation on the top half of the images. Participants had to judge whether the top half of the image was the same (identical) or different (in any way) across the pairs of images. Participants could respond as soon as the second image appeared, and were given a maximum of 3 seconds to respond.

The two images in each sequential pair were always made from parts of the same individual's face, so that face identity was not a confound in the experiment, but the top or bottom parts could differ in expression. Images in each pair either had the same top and bottom halves (no change condition), the same top half but a different bottom half (bottom change condition) or the same bottom half and a different top half (top change condition). The combination of aligned and misaligned variants of these 3 conditions led to 6 conditions in total. Each of these 6 conditions involved 24 trials. Images for the behavioural experiment were presented using PsychoPy2 (Peirce, 2007).

2.3. Experiment 2

2.3.1. Stimuli and Design

Experiment 2 used a block design fMR-adaptation paradigm. In order to identify face-selective regions for each individual, a localiser scan was conducted prior to the experimental scan. The localiser had 3 stimulus conditions: faces, places, and Fourier phase-scrambled faces. Each localiser scan block lasted 9 seconds and contained 9 images from one of the localiser conditions, with each image being presented for 900 milliseconds and a 100 ms inter-stimulus interval (ISI). Each condition was repeated 4 times. Images used in the localiser scan were presented using Neurobehavioural Systems Presentation 16.3.

For the main fMR-adaptation scan, experiment 2 had 6 stimulus conditions (Figure 1) presented in a block design. The same stimuli were used as in experiment 1 to create blocks in which the same image was repeated throughout the block (no change condition), the top half of each image was unchanged throughout the block but the bottom half differed (bottom change condition), or the top half changed while the bottom half stayed the same (top change condition). The use of aligned and misaligned versions of these 3 types of block led to 6 conditions overall. There were equal numbers of aligned and misaligned blocks, and the positioning of the image parts in the misaligned blocks was counterbalanced so that half were misaligned to the left, and half to the right. There were 48 blocks in total (6 conditions, repeated 8 times). For the 8 repetitions of each condition, there were 4 blocks for each of the identities used. Within these 4 blocks, each expression was used once as the top half. This meant that within each condition, each identity and expression combination was presented once.

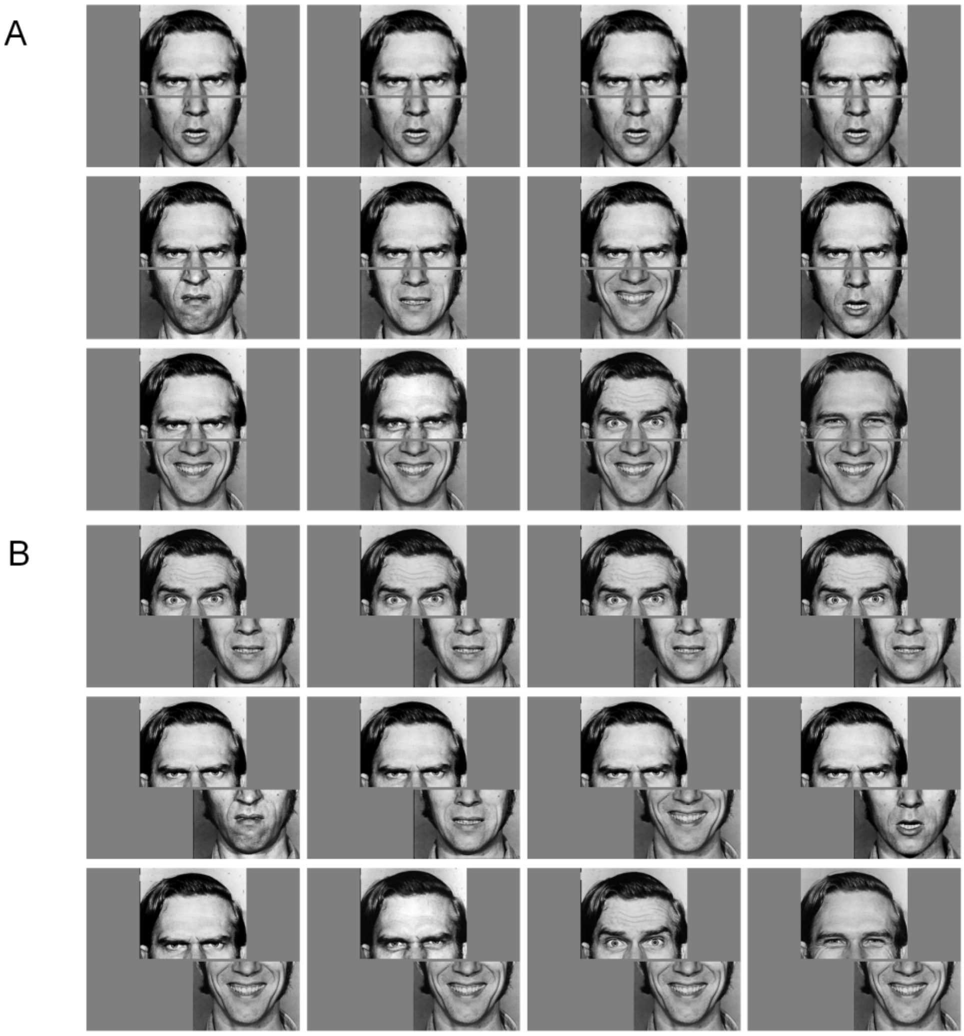

Figure 1. Examples of experimental stimuli used to create trial blocks in experiment 2. A) Aligned conditions (top row: no change, middle row: bottom change, bottom row: top change); B) Misaligned conditions (top row: no change, middle row: bottom change, bottom row: top change). The stimuli used in experiment 1 involved sequentially presented pairs of images from each of the 6 types of trial block. Note that a small gap between the top and bottom halves of each stimulus emphasises where the parts are joined, even for the aligned images (cf. Rossion, 2013).

All images were back-projected onto a screen inside the bore of the scanner, approximately 57cm from the participants' eyes. Images were presented in 6 second blocks; this overall block duration is equivalent to those used in our other recent studies of neural responses to facial expression (Mattavelli et al., 2014; Psalta et al., 2014). Each block contained 4 images, with each image being presented for 750ms with a 750 ms grey screen ISI. There was a 9 second grey screen between each of the blocks. Each stimulus condition was repeated 8 times to give a total of 48 blocks. Hence each scan lasted 12 minutes in total. Images within a block were all derived from the same identity, and the use of each of the 2 identities (models) was randomised across the experiment. Participants monitored all images for the presence of a small red dot (6 pixels in width) that was superimposed at the fixation point on 1 image in each block. Participants were required to respond, with a button press, as soon as they saw the image containing the target red dot. Images for the experimental scan were presented using PsychoPy2 (Peirce, 2007).

2.3.2. Imaging Parameters

All scans were conducted using a GE Signa HDx 3T MRI system (General Electric, Waukesha, WI, USA) with an eight channel phased array head coil (MRI Devices Corp., Gainesville, FL). Data were acquired using a gradient echo planar imaging (EPI) sequence with acquisition parameters: 38 contiguous axial slices, repetition time (TR) 3 seconds, echo time (TE) 32.5 milliseconds, flip angle 90°. The field of view (FOV) was 28.8 x 28.8 cm with an acquisition matrix of 128 x 128 and slice-thickness of 3mm, giving a voxel size of 2.25 x 2.25 x 3mm. A T1-weighted Fluid-Attenuated Inversion Recovery (T1-FLAIR) volume was acquired with the same slice orientation and slice thickness with an acquisition matrix of 512x512, giving an in-plane resolution of 0.5625x0.5625mm. To improve registration, the EPI image was initially co-registered with the high resolution initial structural image (T1-weighted FLAIR) containing the same number of slices as the EPI scan before being registered to the high resolution main structural scan (T1-weighted, 1.13 x 1.13 x 1 mm) for each participant. This was then co-registered to the standard MNI 152 brain.

2.3.3. fMRI Analysis

Analysis was conducted using FEAT v 5.98 (http://www.fmrib.ox.ac.uk/fsl). The initial 9 seconds of each scan were removed from the analysis to allow T1-saturation effects to subside. Motion correction (McFLIRT; FSL) was applied followed by spatial smoothing (Gaussian, Full Width at Half Maximum 6 mm) and temporal high-pass filtering with a cut off of 0.01 Hz. Face-selective regions were defined in each individual from the functional localiser by using the average of the face > place and face > scrambled face contrasts. The combined statistical maps were thresholded at $p < .01$ (uncorrected). For each individual, the OFA, FFA and pSTS were identified by contiguous clusters of voxels activated above threshold from the above contrast in posterior occipital cortex, inferior fusiform gyrus and superior temporal lobe.

For each individual, the time series of the filtered MR data for each voxel from the experimental scan within each functionally localised ROI was converted to percentage signal change. These were then averaged to produce the time series for each participant within each ROI for each of the experimental conditions. The individual time series data were normalised by subtracting each time point by the zero point at the beginning of the block. These data were then averaged across participants to give the overall mean time series for each condition. The peak response to each condition was taken as the average of TR 2 and TR 3 (corresponding to 6 and 9 seconds after stimulus onset). These peak responses were then entered into repeated measures ANOVAs to determine significant differences between conditions for each ROI.

Our primary focus of interest was in neural responses from pSTS based on a functional localiser applied at the individual participant level. However, to determine whether other regions might demonstrate a holistic response to expressions, we also performed a whole brain analysis in which the behavioural data from Experiment 1 were used as regressors. A box car function was defined modelling all blocks in the scan run, with each block weighted by the mean RT of that condition. This was convolved with a single gamma hemodynamic response function and then regressed against the BOLD response at each voxel. The resulting statistical maps for each individual were combined using a higher-level mixed effects analysis (FLAME, FSL). The combined statistical maps were thresholded at $z > 2.8$, $p < .05$ (cluster corrected). This process was then repeated using the % error data as a regressor.

3. RESULTS

3.1. Experiment 1

The aim of experiment 1 was to demonstrate the facial expression composite effect with the stimuli and presentation times to be used in the fMR-adaptation study. There were 6 conditions involving aligned or misaligned pairs with no change between the images, a bottom half change, or a top half change. Participants monitored the top half of pairs of face images to detect whether the facial expression in the top half remained the same, or was different across the two faces.

First we measured the accuracy of responses when judging whether the top half of each image was the same or different. As participants were asked to make their judgements based only the top half of each image, the correct responses in each condition were 'same' for no change pairs, 'same' for the bottom change pairs, and 'different' for the top change pairs. Percent correct responses were calculated for each condition for each participant, and then averaged across all participants to give an overall percent correct response measure. The data are displayed as percentage errors in Figure 2A to facilitate comparison with reaction times shown in Figure 2B.

The proportion of correct responses was entered into a 2 x 3 repeated measures ANOVA with the factors Alignment (aligned, misaligned) and Condition (no change, top change, bottom change). The ANOVA showed a significant effect of Alignment ($F(1,15) = 38.37$, $p < .001$, partial eta squared = 0.72) and Condition ($F(2,30) = 19.48$, $p < .001$, partial eta squared = 0.57). Bonferroni pairwise comparisons demonstrated that the effect of Alignment was driven by more accurate responses in the misaligned versus aligned conditions ($p < .001$). The effect of Condition was driven by more accurate responses in the no change versus top change ($p < .001$) and bottom change ($p = .001$) conditions. However, these main effects were both qualified by the presence of a significant Alignment x Condition interaction ($F(2,30) = 10.82$, $p < .001$, partial eta squared = 0.42). Paired t-tests demonstrated this was a result of lower accuracy in the bottom change condition when the stimuli were aligned, compared to misaligned ($t(15) = -5.54$, $p < .001$) but no difference between the no change aligned and misaligned conditions ($t(15) = .432$, $p = .672$). This part of the interaction is the critical test of the facial composite effect, because in all four of these conditions participants were making equivalent responses (that the top halves were the 'same'). In addition, there was also a non-significant trend demonstrating lower accuracy for the top change condition when the stimuli were aligned, compared to misaligned ($t(15) = -1.86$, $p = .083$). Whilst of interest, this is less crucial because the correct response has now switched to 'different'.

We also measured response times to each condition. Median RTs were taken for each condition, for each participant and an overall median RT was calculated for each condition across all participants (Figure 2B). These median RTs were entered into a 2 x 3 repeated measures ANOVA with the factors Alignment (aligned, misaligned) and Condition (no change, top change, bottom change). This ANOVA demonstrated significant main effects of Alignment ($F(1,15) = 18.24$, $p = .001$, partial eta squared = 0.55) and Condition ($F(2,30) = 16.36$, $p < .001$, partial eta squared = 0.52). Bonferroni pairwise comparisons demonstrated the effect of Alignment was driven by longer RTs when the stimuli were aligned, compared to misaligned ($p = .001$) and the effect of Condition was driven by a longer RT in both top change ($p = < .001$) and bottom change ($p < .001$) conditions relative to no change.

Again, interpretation of these main effects needs to be qualified by a significant Alignment x Condition interaction ($F(2,30) = 11.62$, $p < .001$, partial eta squared = 0.44). Paired t-tests demonstrated this was due to longer response times in the aligned versions of both top change and bottom change conditions when compared to their misaligned counterparts (bottom change: $t(15) = 4.69$, $p < .001$, top change: $t(15) = 3.04$, $p < .001$). No difference was seen in the response times between the aligned and misaligned versions of the no change condition ($t(15) = -1.54$, $p = .145$). Paralleling the analysis of accuracy data, the slower response times in the aligned compared to misaligned version of the bottom change condition, and the lack of difference in response time for the no change condition, illustrate the key components of the face composite effect.

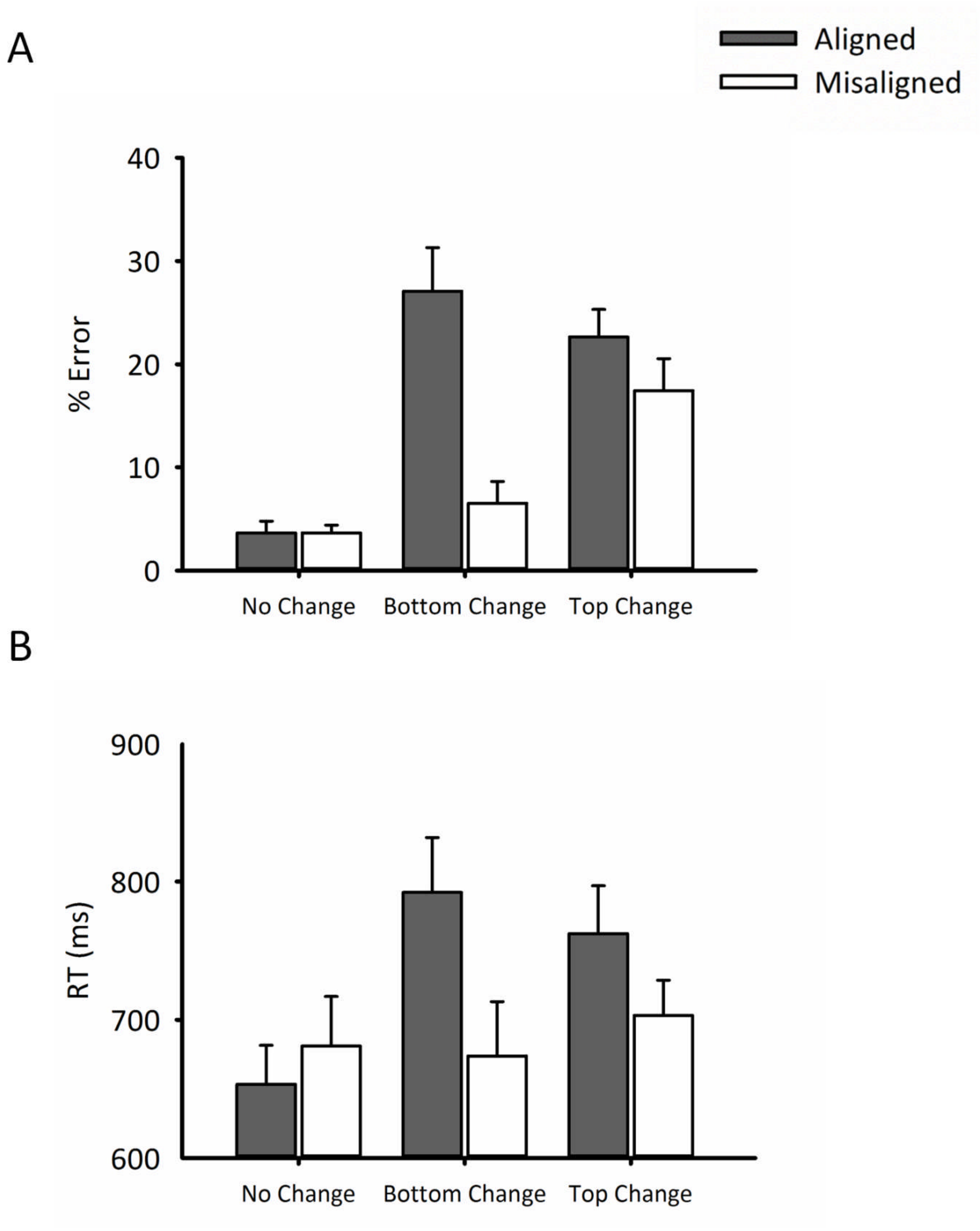

Figure 2. (A) Percent error responses for the same-different task in Experiment 1. The critical result is reduced performance (increased errors) in the bottom change condition compared to the no change condition when the stimuli are aligned, but not when they are misaligned, even though in all 4 of these conditions the top halves of the stimuli are to be judged 'same' – the facial expression composite effect. The top change condition is less important because it involves a change in correct response (now 'different' instead of 'same'). (B) Median response times for the same-different task in experiment 1. RTs were longer for the bottom change condition compared to no change condition when the stimuli were aligned, but not misaligned – again demonstrating the expression composite effect.

In sum, behavioural results from the RT and accuracy data show the facial expression composite effect where participants find it more difficult to judge the top half of the images as the same when the bottom half is changing and the two halves of each image are aligned into an overall facial configuration, compared to when they are in a misaligned form.

3.2. Experiment 2

The aim of this experiment was to investigate properties of the right pSTS response to facial expressions, using conditions comparable to those in the behavioural experiment 1. The principal focus of the analysis was pSTS because of its hypothesised role in facial expression perception in the leading neural model of face perception, (Haxby et al., 2000), and on right rather than left pSTS because right pSTS is more reliably identified at the individual participant level with our functional localiser scan and has therefore been targeted in previous studies (Harris et al., 2012, 2014). To parallel experiment 1, there were 6 different types of block in the experimental scan, involving aligned or misaligned pairs with no change between the images, a bottom half change, or a top half change

In order to check whether participants were watching the top halves of the stimuli throughout the experiment, as instructed, they were given the task of pressing a response button every time they saw a small red dot presented at the fixation point. Performance on this red dot detection task was high, with a mean accuracy of 99% correct responses and mean RT of 447ms. To confirm that there were no differences in overall attentional demands between aligned and misaligned stimuli, the average response times to aligned and misaligned conditions for each participant were entered into a paired t-test. There was no significant difference in response times to the red dot, $t(21)=1.39$, $p = .18$.

The pSTS, FFA and OFA were localised in the left and right hemispheres using the independent functional localiser scan. The OFA and FFA could be identified in both the left and right hemispheres for 23/26 participants. In contrast to the OFA and FFA, the pSTS could be reliably identified in the right hemisphere of 22/26 participants, but in the left hemisphere for only 15/26 participants. This relatively poor face responsiveness of left pSTS may be due to its possible role in more audiovisual integration of vocal and facial speech signals (Calvert, 2001; Pelphrey et al., 2005; Wright et al., 2003). Average MNI coordinates and number of voxels for each localised ROI are provided in Table 1.

Table 1. *Average MNI coordinates in mm (mean and SE), size in voxels, and number of participants where the region could be identified, for each ROI.*

ROI	Coordinate			No. of Voxels	No. of Participants
	x	y	z		
Right OFA	41 ± 1	-80 ± 2	-15 ± 1	187	26
Left OFA	-41 ± 1	-83 ± 1	-14 ± 1	107	23
Right FFA	41 ± 1	-56 ± 1	-23 ± 1	223	26
Left FFA	-40 ± 1	-60 ± 2	-23 ± 1	114	23
Right pSTS	51 ± 1	-61 ± 2	1 ± 1	110	23

There was no effect of hemisphere for the OFA ($F(1,22) = 0.16$, $p = .696$) or FFA ($F(1,22) = 1.58$, $p = .221$), so the data from the left and right hemispheres of these regions were combined. For pSTS, we used only the region localised in the right hemisphere. In terms of Haxby et al.'s (2000) neural model of face perception, results for the pSTS and FFA are the most instructive, as these lie on separate neural pathways considered to be critically involved in the perception of expression (pSTS) or to be involved in other aspects of face perception (FFA). Data for the pSTS and FFA are therefore summarised in Figure 3. The OFA was considered as of less interest because it lies on both neural pathways in Haxby et al.'s (2000) model, but data from the OFA were analysed, for completeness.

First, we took the time series data for each participant and averaged these across participants to give an overall mean time series for each condition, for each ROI (Figure 3). We then looked at the peak responses in the right pSTS, which form the study's principal focus of interest (Figure 3, panel A). A 2x3 ANOVA with the factors Alignment (aligned, misaligned) and Condition (no change, bottom change, top change) demonstrated a significant effect of Condition ($F(2,44) = 7.62$, $p = .001$), but not of Alignment ($F(1,22) < 1$). The Alignment x Condition interaction was not significant ($F(2,44) < 1$). The effect of Condition was driven by a smaller peak percentage signal change in the no change condition compared to both the bottom change ($t(22) = -3.75$, $p = .001$) and top change conditions ($t(22) = -2.93$, $p = .008$), with no difference between the signal change in the bottom and top change conditions ($t(22) = .301$, $p = .797$). This pattern is consistent with a feature-based response, with no evidence of the critical interaction between Alignment and Condition that would demonstrate holistic perception.

It is important to note that in this study, we looked at the response across all facial expressions. Although our design does not allow for the data to be explored in this way, it would be interesting to look at the response for each individual expression. This would be particularly interesting as some facial expressions are more recognisable from their bottom halves, and some from their top halves (Calder et al. 2000).

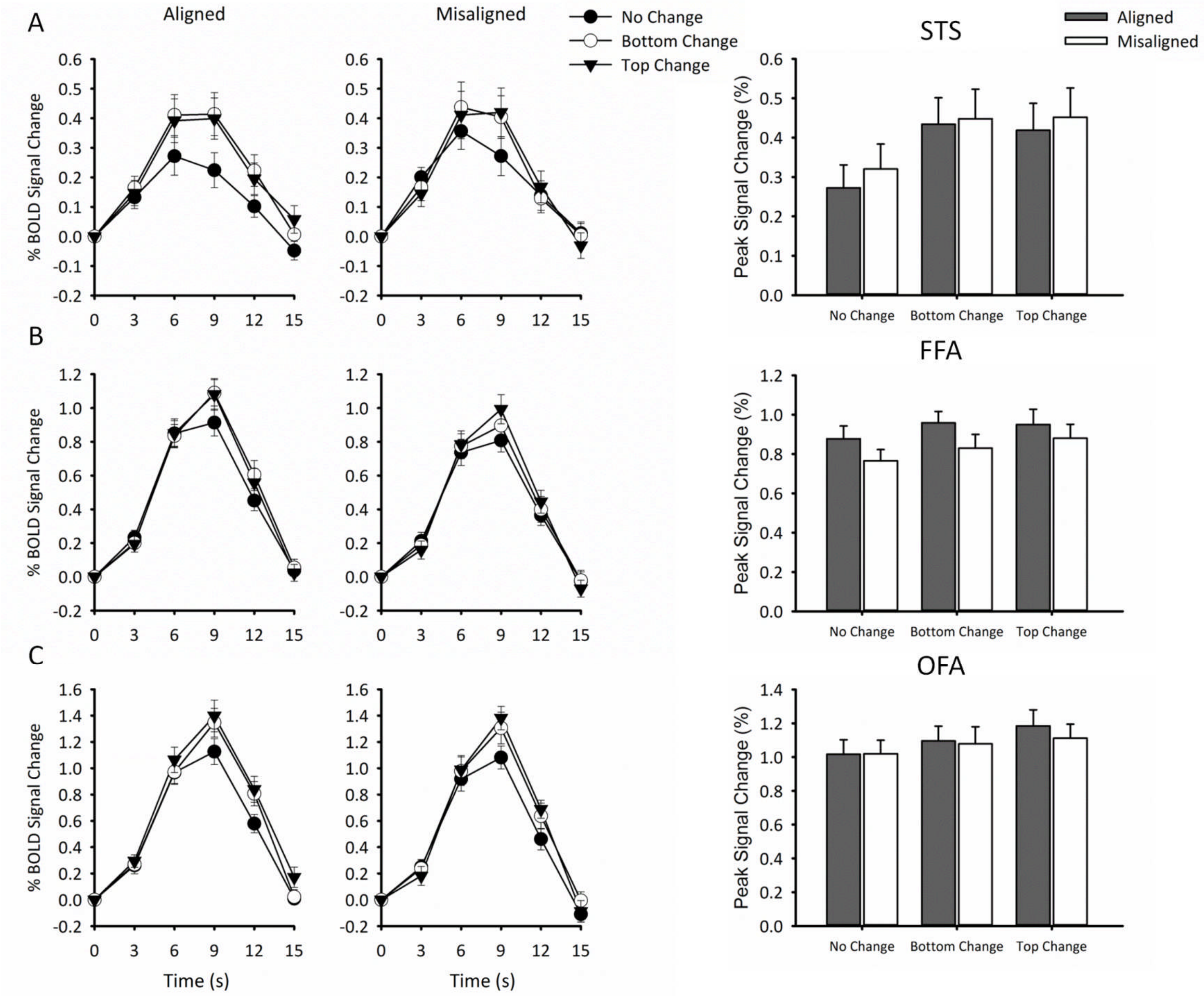

Figure 3. Overall mean MR time series for each condition for aligned and misaligned stimuli, and peak % BOLD signal change for right pSTS (row A), FFA (row B) and OFA (row C). Analysis of the responses in right pSTS revealed a smaller peak response in the no change condition compared to both the bottom change ($p = .001$) and top change conditions ($p = .008$), with no difference between the bottom and top change conditions. This pattern held for aligned and misaligned stimuli. In FFA, there was only a main effect of Alignment, with a higher peak response to aligned than misaligned stimuli ($p = .021$). Error bars represent standard error of the mean.

The FFA showed a different pattern of results to the pSTS (Figure 3, panel B). A 2x3 ANOVA showed a significant effect of Alignment ($F(1,25) = 6.11$, $p = .021$), but only a borderline effect of Condition ($F(2,50) = 2.56$, $p = .088$). The Alignment x Condition interaction was not significant ($F(2,50) < 1$. The effect of Alignment was driven by a significantly higher peak percent signal change to the aligned compared to misaligned stimuli ($t(25) = 2.47$, $p = .021$).

The OFA did not produce any findings that reached conventional levels of statistical significance (Figure 3). There was no effect of Alignment ($F(1,25) < 1$, and after Greenhouse-Geisser correction for a violation of sphericity ($\chi^2(2) = 9.03$, $p = .011$) only a borderline effect of Condition ($F(1.523,38.07) = 3.32$, $p = .059$). There was no Alignment x Condition interaction ($F(2,50) < 1$).

To determine if other regions showed a holistic response, we also conducted a whole brain analysis. The % error and response time data from Experiment 1 were used as regressors to identify regions that might show a holistic response. The resulting group statistical parametric map identified 2 clusters of activity, in the right inferior frontal gyrus (IFG) and in the right fusiform gyrus. Table 2 shows the peak voxel intensity, co-ordinates and size of the ROIs based on the % error and RT data.

Table 2. *Peak intensity and MNI coordinates (mm) for maximally active voxel, and size in voxels for each ROI identified using the mean RT and % error data from experiment 1 as a regressor.*

ROI	Peak Intensity (z score)	Coordinate			No. of Voxels
		x	y	z	
% Error					
Right Fusiform	4.86	38	-50	-22	771
Right IFG	3.90	48	4	18	411
RT					
Right Fusiform	4.97	40	-50	-24	656
Right IFG	4.09	48	6	18	654

These data were used to create masks of the regions identified (right fusiform, and right IFG). We took the time series data for each participant and averaged across participants to give an overall mean time series for each condition, for each ROI. The peak responses for each condition for each ROI were then calculated. As can be seen from table 2, the peak intensities were very similar for both the ROIs identified using the RT and % error data. This was also reflected in the peak response to each individual condition, therefore we have only presented the % error regressor data for illustration purposes, in Figure 4. The right IFG shows the classic pattern demonstrated in the expression composite effect – a higher response to bottom change when the face is aligned, compared to when misaligned. It also shows a smaller response to the no change compared to the change conditions. In contrast, the fusiform gyrus shows a more general overall difference in responsiveness between aligned and misaligned images. This is consistent with the known involvement of fusiform cortex in the holistic perception of faces (Kanwisher et al., 1997; Andrews et al., 2010), but does not imply holistic processing of expression per se.

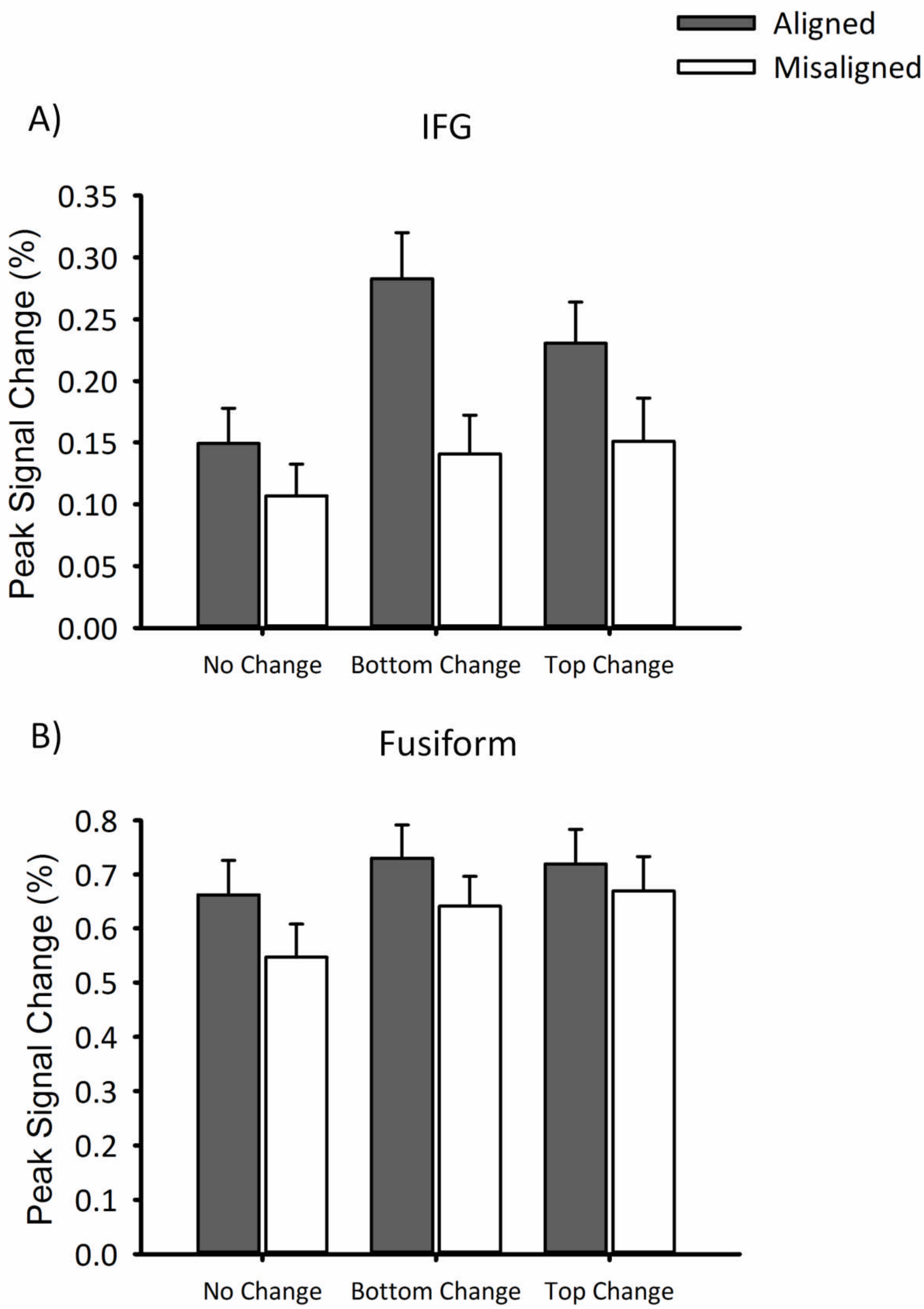

Figure 4. Overall mean peak % BOLD signal change for each condition for aligned and misaligned stimuli for the right IFG (A), and right fusiform (B). Regions defined using the % error data from experiment 1 as a regressor. Error bars represent standard error of the mean.

4. DISCUSSION

We used the fMR-adaptation paradigm to investigate neural responses to facial expressions in core regions of Haxby et al.'s (2000) neural model of face perception, focussing particularly on pSTS because of its hypothesised role in the perception of expression. By using a no change condition as a baseline promoting maximal adaptation, we were able to demonstrate release from adaptation in right pSTS to conditions in which changes in expression were located in the upper or lower parts of the stimuli. This shows that the right pSTS was encoding such changes, even though the incidental task of detecting a red spot was irrelevant to perceiving the facial expression. Moreover, the degree of adaptation in right pSTS was equivalent whether the changes occurred in the fixated, task-relevant (top half) or non-fixated (bottom half) part of each stimulus.

This pattern of neural response in pSTS was the same regardless of whether the top and bottom parts of the stimuli were aligned into a face-like overall configuration, or misaligned by offsetting the parts to make the overall image less face-like. The contrast between aligned and misaligned variants of the stimuli is of theoretical importance, as it is now widely used to probe holistic processing of faces in studies of the perception of face identity and facial expression (Young et al., 1987; Calder et al., 2000; Maurer et al., 2002; Rossion, 2013). The logic underlying the contrast is that holistic processing of the stimulus as a face is only possible when the constituent parts are correctly aligned, and that a consequence of holistic processing will be to enhance perceived differences between stimuli that share common parts - for example, making the top change stimuli look more different from each other when in the aligned than in the misaligned arrangement. This enhanced perception of differences between aligned than misaligned stimuli was demonstrated behaviourally in Experiment 1, so it is noteworthy that our results do not show such an effect in the neural responses from pSTS. Instead, it seems that pSTS is sensitive to any change in face parts (with a release from adaptation in both top change and bottom change conditions) but does not require that the stimulus is particularly face-like (as shown by the equivalent release from adaptation across aligned and misaligned stimuli). This complements Harris et al.'s (2012) finding that pSTS responds more or less linearly to all changes in facial features that communicate emotion.

A possibility that needs to be considered is that the differences in the pattern of the results between the behavioural (Experiment 1) and fMRI (Experiment 2) data might reflect task differences. In the behavioural experiment, participants were asked to detect changes in facial expression. In contrast, in the fMRI experiment, participants were asked to detect a red dot superimposed on some of the faces. An explicit holistic task was not used in the fMRI experiment because our aim was to examine how facial expression is encoded irrespective of task difficulty. Using an explicit task of holistic processing would introduce differences in task difficulty across conditions and as a result, produce attentional differences across conditions. Therefore it was important to use a task independent of the experimental manipulation to ensure all stimuli were attended to equally in the fMR experiment. Since the expression composite effect is considered to reflect mandatory holistic face perception and no previous work has suggested that it is affected by the task, this offered the best way to eliminate potential attentional confounds. It is also important to note that the facial *identity* composite effect can be demonstrated using a similar fMRI experimental procedure (Schiltz and Rossion, 2006) to that presented here.

The FFA showed a different pattern of response than pSTS, with the only finding that reached the conventional level of statistical significance being a main effect of alignment, with higher overall response to aligned than to misaligned stimuli. These results are consistent with previous studies that used fMR-adaptation with composite faces to reveal a holistic response to facial identity in the FFA (Schiltz and Rossion, 2006; Schiltz et al., 2010; Andrews et al., 2010). The pattern is also consistent with Kanwisher et al.'s (1997) landmark study defining the properties of the FFA, which found a stronger response to normal faces than to scrambled arrangements of face parts, as misaligning the stimuli can be considered a simple variant of face scrambling. This finding reveals that there are fundamentally different neural representations of faces in the FFA and pSTS. The representation in the FFA is sensitive to the correct configuration of the facial features, whereas the pSTS appears to encode facial features independently.

To determine if regions outside the core face-selective regions showed a holistic response to facial expression, we performed a group analysis. This analysis used the behavioural data from Experiment 1 as a regressor, as this had shown a holistic response to expression. The independence of the behavioural (Experiment 1) and fMRI (Experiment 2) data used in this analysis offers a strong test of whether a region can be linked to a specific pattern of responses. This group analysis identified the right fusiform gyrus and right inferior frontal gyrus (IFG) as regions that covaried with behavioural responses. Inspection of the data shown in Figure 4 suggests that the fusiform activity was due to a more general holistic response to faces per se, in the form of a higher overall response to all aligned than misaligned stimuli, as had also been shown from the analysis of the FFA defined with the individually-based functional localiser. In contrast, the IFG showed a pattern of response which was more consistent with a holistic response to facial expression, as evidenced by the similarity between the pattern of BOLD responses in IFG (Figure 4) and the RTs and errors in the behavioural task (Figure 2). These results are consistent with previous studies which have shown that right IFG is part of the extended face processing network (Ishai et al. 2008; Davies-Thompson et al., 2012) and is involved in the processing of facial expressions (Ishai, Schmidt & Boesiger, 2005; Carr et al. 2003; Dapretto et al., 2006).

In sum, we have shown that right pSTS is sensitive to changes in the facial features that convey emotion regardless of whether these changes occur in the fixated parts of the image or not, and regardless of whether image parts are arranged in a more or a less face-like configuration. Therefore, based on these results, the pSTS cannot be considered the neural locus of the facial expression composite effect. Nonetheless, these findings are consistent with Haxby et al.'s (2000) view that pSTS is an important region in the perceptual analysis of facial expressions and uncover something of this region's *modus operandi*, showing in particular that it is very responsive to changes in expressive features whether or not these form a face-like overall configuration.

Acknowledgements

TRF was supported by a studentship from the University of York.

REFERENCES

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278.

Andrews, T.J., Davies-Thompson, J., Kingstone, A., & Young, A.W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience*, *30*, 3544-3552.

Baseler, H.A., Harris, R.J., Young, A.W., & Andrews, T.J. (2014). Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cerebral Cortex, 24*, 737-744.

Bruce, V. & Young, A.W. (1998). *In the eye of the beholder: the science of face perception*. Oxford: Oxford University Press.

Bruce, V. & Young, A.W. (2012). *Face perception*. Hove, East Sussex: Psychology Press.

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. *Visual Cognition, 12*, 495–518.

Calder, A.J. & Young, A.W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 641-651.

Calder, A.J., Young, A.W., Keane, J., & Dean, M. (2000). Configurational information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 527-551.

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, *11*, 1110–1123.

Carr, L., Iacoboni, M., Dubeau, M-C., Mazziotta., J.C. & Lenzi, G.L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, USA, 100*(9), 5497-5502.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. a, Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9, 28-30.

Davies-Thompson J & Andrews TJ (2012). Intra- and interhemispheric connectivity between face-selective regions in the human brain. *Journal of Neurophysiology, 108*, 3087-3095.

Harris, R.J., Young, A.W., & Andrews, T.J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proceedings of the National Academy of Sciences, USA, 109*, 21164-21169.

Harris, R.J., Young, A.W., & Andrews, T.J. (2014). Brain regions involved in processing facial identity and expression are differentially selective for surface and edge information. *NeuroImage, 97*, 217-223.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Ishai, A. (2008) Let's face it: It's a cortical network. *NeuroImage, 40*(2), 415-419.

Ishai, A., Schmidt, C.F. & Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin, 67,* 87-93.

Johnston, P., Mayes, A., Hughes, M., & Young, A.W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex, 49*, 2462-2472.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302-4311.

Mattavelli G., Sormaz M., Flack T., Asghar A.U.R., Fan S., Frey J., Manssuer L., Usten D., & Young A.W., Andrews T.J. (2014). Neural responses to facial expressions support the role of the amygdala in processing threat. *Social Cognitive and Affective Neuroscience, 9* (11), 1684-1689.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*, 255-260.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia,49*, 1226–1235.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8–13.

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, *15*, 1866 – 1876.

Prazak, E.R. and Burgund, E.D. (2014). Keeping it real: Recognizing expressions in real compared to schematic faces. *Visual Cognition, 22*(5), 737-750.

Psalta, L., Young, A.W., Thompson, P. and Andrews, T.J. (2014). The Thatcher illusion reveals orientation-dependence in brain regions involved in processing facial expression. *Psychological Science*, *25*(1), 128-136.

Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic face perception. *Visual Cognition*, *21*, 139-253.

Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision, 10*, 1-12.

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. Journal of Vision, 10(2), 1-16.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipitotemporal cortex. *NeuroImage, 32*, 1385-1394.

Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex*, *13*, 1034 – 1043.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747-759.

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., Ekman, P. (2002). *Facial expressions of emotion stimuli and tests (FEEST)*. Bury St. Edmunds, Suffolk, England: Thames Valley Test Company.